

J.-J. DAUDIN

Étude de l'amplitude d'une marche aléatoire de Bernoulli

RAIRO. Recherche opérationnelle, tome 30, n° 1 (1996),
p. 99-106

http://www.numdam.org/item?id=RO_1996__30_1_99_0

© AFCET, 1996, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ÉTUDE DE L'AMPLITUDE D'UNE MARCHE ALÉATOIRE DE BERNOULLI (*)

par J.-J. DAUDIN ⁽¹⁾

Communiqué par Jean ABADIE

Résumé. – Vallois et Tapiero (1994) ont étudié les deux premiers moments du processus d'amplitude d'une marche aléatoire. Dans cette Note, nous obtenons les mêmes résultats par une méthode plus directe. Cette démonstration permet de plus d'obtenir de nouvelles propriétés comme les moments centrés d'ordre 3 et 4, ainsi qu'une relation de récurrence permettant d'obtenir la fonction génératrice.

Mots clés : marche binomiale symétrique, amplitude, moments, fonction génératrice.

Abstract. – The range process for a symmetric Bernoulli random walk has been studied by Vallois and Tapiero. They obtained the first two moments of the inverse process. In this Note, we obtain the same result by a more direct way. Moreover the third and fourth centered moments are obtained by the same way and a recurrent relation is given for the probability generating function.

Keywords: range process, symmetric bernoulli random walk, moments, moment-generating function.

DÉFINITION 1 : $X_1 \dots X_t$ est un processus de Bernoulli sans biais :

$X_t = X_{t-1} + E_t$ où E_t est une variable aléatoire prenant les valeurs 1 et -1 avec les probabilités suivantes : $P(E_t = 1) = P(E_t = -1) = 0,5$. Les variables E_t sont indépendantes et $X_0 = 0$.

On définit l'amplitude, A_t , du processus par $A_t = \max[X_1 \dots X_t] - \min[X_1 \dots X_t]$.

$\theta_a = \inf\{t > 0 / A_t = a\}$ est l'inverse du processus d'amplitude, c'est-à-dire le temps minimum pour atteindre l'amplitude a .

(*) Reçu en avril 1994.

(1) Département de Mathématique et Informatique de l'INAPG, Laboratoire associé de Biométrie de l'INRA. 16, rue Claude Bernard, 75231 Paris Cedex 05.

THÉOREME 1 : *L'espérance, la variance et les moments centrés d'ordre 3 et 4 de θ_a sont donnés par :*

$$\begin{aligned} E(\theta_a) &= [a(a+1)]/2 \\ V(\theta_a) &= [a(a-1)(a+1)(a+2)]/12 \\ E[(\theta_a - E(\theta_a))^3] &= [a(a-1)(a+1)(a+2)(2a^2 + 2a + 3)]/60 \\ E[(\theta_a - E(\theta_a))^4] &= [a(a-1)(a+1)(a+2)(69a^4 + 138a^3 + 115a^2 + 46a - 100)]/1680 \end{aligned}$$

COROLLAIRE 1 : *Le coefficient d'asymétrie de θ_a est égal à*

$$As(\theta_a) = 2\sqrt{3}(2a^2 + 2a + 3)/5[a(a-1)(a+1)(a+2)]^{1/2}$$

$\lim As(\theta_a) = 4\sqrt{3/5} \cong 1.3856$ *quand a tend vers l'infini.*

Le coefficient d'aplatissement de θ_a est égal à

$$Ap(\theta_a) = (3/35)(69a^4 + 138a^3 + 115a^2 + 46a - 100)/[a(a-1)(a+1)(a+2)]$$

$\lim Ap(\theta_a) = 207/35 \cong 5.9143$ *quand a tend vers l'infini.*

La démonstration du théorème utilise la définition 2, les lemmes 1 et 2 et leurs corollaires ainsi que la proposition 1, qui est le résultat essentiel de ce papier.

DÉFINITION 2 : Soit Y_t un processus de Bernoulli sans biais avec $Y_0 = 0$. Les variables aléatoires t_a et u_a sont définies par les relations suivantes pour $a \geq 2$:

$$\begin{aligned} t_a &= \inf\{t > 0 / [(Y_t = 1) \cup (Y_t = -a)]\} \\ u_a &= \inf\{t > 0 / [(Y_t = -1) \cup (Y_t = a)]\} \end{aligned}$$

LEMME 1 : *Le temps de passage de l'amplitude $a-1$ à l'amplitude a est distribué comme t_a ou u_a . Plus précisément on a :*

si $X_{\theta_{a-1}} = \max[X_1 \dots X_{\theta_{a-1}}]$, alors $\theta_a - \theta_{a-1} = t_a$

si $X_{\theta_{a-1}} = \min[X_1 \dots X_{\theta_{a-1}}]$, alors $\theta_a - \theta_{a-1} = u_a$

où $Y_t = X_{\theta_{a-1}+t} - X_{\theta_{a-1}}$

Démonstration du lemme 1 : L'amplitude $a-1$ est atteinte au temps θ_{a-1} pour $X_{\theta_{a-1}} = \max[X_1 \dots X_{\theta_{a-1}}]$ ou pour $X_{\theta_{a-1}} = \min[X_1 \dots X_{\theta_{a-1}}]$.

Supposons que l'amplitude $a - 1$ soit atteinte au temps θ_{a-1} pour $X_{\theta_{a-1}} = \max[X_1 \dots X_{\theta_{a-1}}]$; soit Y_h le processus défini par $Y_h = X_{\theta_{a-1}+h} - X_{\theta_{a-1}}$; pour que X_t atteigne l'amplitude a , il faut et il suffit que Y_h atteigne la valeur 1 ou la valeur $-a$. Or Y_h est un processus de Bernoulli sans biais avec $Y_0 = 0$. t_a est donc le temps de passage de l'amplitude $a - 1$ à l'amplitude a si $X_{\theta_{a-1}} = \max[X_1 \dots X_{\theta_{a-1}}]$.

Le même raisonnement vaut pour u_a . ♦

LEMME 2 : t_a et u_a ont la même distribution de probabilité.

Le lemme 2 est dû au fait que le processus est sans biais. Si $P(E_t = 1) = p$ avec $p \neq 0.5$, on perd cette propriété.

COROLLAIRE 2 : θ_a et $\sum_{i=1,a} t_i$ ont même loi de probabilité.

COROLLAIRE 3 : Soit $g_a(s)$ la fonction génératrice de θ_a et $h_i(s)$ celle de t_i . On a $g_a(s) = \prod_{i=1,a} h_i(s)$.

Le corollaire est dû au fait que les t_i sont des variables aléatoires indépendantes, car ce sont des temps d'arrêt sur des périodes différentes du processus.

La proposition suivante (proposition 1) est l'élément clé de la démonstration des théorèmes 1 et 2.

PROPOSITION 1 : Soit ε une variable de Bernoulli indépendante des t_i , prenant les valeurs 0 et 1 avec la probabilité 1/2. On pose $t_0 = 0$ et $t_1 = 0$. Pour $i \geq 2$, les variables aléatoires t_i et $1 + \varepsilon(t_{i-2} + t'_i)$ ont la même loi de probabilité, où t'_i et t_{i-2} sont indépendantes et t_i et t'_i ont même loi.

Démonstration de la proposition : $t_i = \inf\{t > 0 / [(X_t = 1) \cup (X_t = -i)]\}$ avec $X_0 = 0$.

Deux cas se présentent :

- 1) $X_1 = 1$; dans ce cas $t_i = 1$.
- 2) $X_1 = -1$; dans ce cas le processus doit d'abord atteindre l'une des barrières à 0 ou à $-i + 1$ avant d'atteindre 1 ou $-i$. On a donc :

$$t_i = 1 + u + v$$

où :

$$u = \inf\{t > 0 / [(X_{t+1} - X_1 = 1) \cup (X_{t+1} - X_1 = -i + 2) / X_1 = -1]\}$$

$$v = \inf\{t > 0 / [(X_{t+u+1} - X_{u+1} = 1) \cup (X_{t+u+1} - X_{u+1} = -i) / (X_{u+1} - X_1 = 1)] \quad \text{si } X_{u+1} - X_1 = 1$$

$$v = \inf\{t > 0 / [(X_{t+u+1} - X_{u+1} = -1) \cup (X_{t+u+1} - X_{u+1} = i) / (X_{u+1} - X_1 = 1)] \quad \text{si } X_{u+1} - X_1 = -i + 2$$

Comme u a la même loi que t_{i-2} et v a la même loi que t_i ou u_i on a donc finalement :

$$t_i = 1 + [(1 - X_1)/2](t_{i-2} + t'_i) \quad \text{ou} \quad t_i = 1 + [(1 - X_1)/2](t_{i-2} + u_i)$$

Comme t'_i et u_i ont la même loi de probabilité, la proposition 1 est démontrée. ♦

On peut illustrer cette démonstration par la figure suivante montrant les trajectoires du processus :

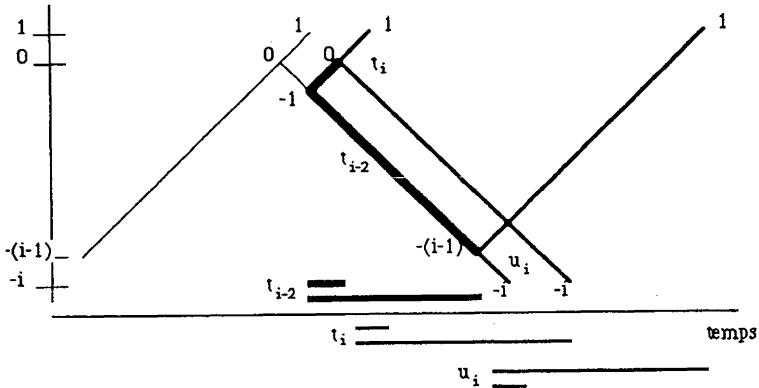


Figure. - Illustration de la proposition 1 dans le cas où $X_{\theta_{i-1}} = \max[X_1 \dots X_{\theta_{i-1}}]$.

Démonstration du théorème 1 : D'après la proposition 1, on a :

$$E(t_i) = 1 + [E(t_{i-2}) + E(t_i)]/2$$

avec $E(t_0) = 0$ et $E(t_1) = 1$.

On en déduit que $E(t_i) = i$ et d'après le corollaire 2, $E(\theta_a) = a(a + 1)/2$.

De même d'après la proposition 1 la variable centrée $t_i - i$ a la même loi que :

$$\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)$$

En élevant au carré ces deux variables, en considérant les espérances associées on a :

$$E(t_i - i)^2 = E\{\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)\}^2$$

en utilisant le fait que ε , t_{i-2} et t_i sont indépendantes, on en déduit

$$V(t_i) = V(t_{i-2}) + 2(i-1)^2$$

et donc

$$V(t_i) = i(i^2 - 1)/3$$

Pour obtenir ce dernier résultat, on remarque que la somme des carrés des nombres impairs inférieurs à un nombre pair j est égale à $j(j^2 - 1)/6$, et que la même formule vaut également pour la somme des carrés des nombres pairs inférieure à un nombre impair.

D'après le corollaire 2 et en utilisant le fait que les t_i sont des variables aléatoires indépendantes, on trouve :

$$V(\theta_a) = \sum_{i=1, a} V(t_i) = [a(a-1)(a+1)(a+2)]/12$$

Le moment centré d'ordre 3 se traite de la même façon; d'après la proposition 1 la variable centrée $t_i - i$ a la même loi que :

$$\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)$$

En élevant au cube ces deux variables, en considérant les espérances associées on a

$$E(t_i - i)^3 = E\{\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)\}^3$$

en utilisant le fait que ε , t_{i-2} et t_i sont indépendantes, on en déduit

$$E(t_i - i)^3 = E(t_{i-2} - (i-2))^3 + 2(i-1)^4 + 4(i-1)^2$$

et donc

$$E(t_i - i)^3 = i(i^4 - 1)/5$$

Pour obtenir ce dernier résultat, on remarque que la somme des puissances quatrièmes des nombres impairs inférieurs à un nombre pair j , est égale à $j(j^2 - 1)^2/10 - 2(j(j^2 - 1))/15$, et que la même formule vaut également pour la somme des puissances quatrièmes des nombres pairs inférieure à un nombre impair.

D'après le corollaire 2 et en utilisant le fait que les t_i sont des variables aléatoires indépendantes, on trouve :

$$\begin{aligned} E[(\theta_a - E(\theta_a))^3] &= \sum_{i=1,a} E(t_i - E(t_i))^3 \\ &= [a(a-1)(a+1)(a+2)(2a^2 + 2a + 3)]/60 \end{aligned}$$

Enfin, le moment centré d'ordre 4 se traite de la même façon; d'après la proposition 1 la variable centrée $t_i - i$ a la même loi que :

$$\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)$$

En élevant à la puissance quatrième ces deux variables, et en considérant les espérances associées on a

$$E(t_i - i)^4 = E\{\varepsilon[t_{i-2} - (i-2) + t_i - i] + (i-1)(2\varepsilon - 1)\}^4$$

en utilisant le fait que ε , t_{i-2} et t_i sont indépendantes, on en déduit après une page de calculs simples mais laborieux :

$$\begin{aligned} E(t_i - i)^4 &= E(t_{i-2} - (i-2))^4 \\ &\quad + [2/15]\{17(i-1)^6 + 30(i-1)^5 + 110(i-1)^4 \\ &\quad + 60(i-1)^3 + 68(i-1)^2\} \end{aligned} \quad (1)$$

et donc

$$E(t_i - i)^4 = i(i^2 - 1)(17i^4 + 35i^3 + 52i^2 - 35i - 25)/105 \quad (2)$$

Pour obtenir ce dernier résultat, on a utilisé un logiciel de calcul formel pour faire la somme du polynôme en $i-1$ de l'équation (1), et on remarque une fois encore que la même formule fonctionne pour i pair ou impair.

D'après le corollaire 2 et en utilisant le fait que les t_i sont des variables aléatoires indépendantes, on trouve :

$$\begin{aligned} E[(\theta_a - E(\theta_a))^4] &= \sum_{i=1,a} E(t_i - E(t_i))^4 \\ &\quad + 6 \sum_{k>1} E(t_k - E(t_k))^2 E(t_1 - E(t_1))^2 \end{aligned} \quad (3)$$

Il est facile d'obtenir le premier terme en sommant la relation (2) :

$$\sum_{i=1,a} E(t_i - E(t_i))^4 = a(a^2 - 1)(a + 2)(17a^4 + 74a^3 + 135a^2 + 6a - 86)/840$$

Le deuxième terme s'obtient en remarquant que

$$\begin{aligned} & \sum_{k>1} E(t_k - E(t_k))^2 E(t_1 - E(t_1))^2 \\ &= [1/2] \{ [\sum_{i=1,a} E(t_i - E(t_i))^2]^2 - \sum_{i=1,a} E(t_i - E(t_i))^2 \} \end{aligned}$$

Le premier terme à droite du signe égal n'est autre que $[V(\theta_a)]^2$, et le deuxième est égal à la somme des carrés des variances des t_i . On a donc ;

$$\sum_{i=1,a} [E(t_i - E(t_i))^2]^2 = \sum_{i=1,a} i^2(i^2 - 1)^2/9$$

d'où on déduit

$$\sum_{i=1,a} [E(t_i - E(t_i))^2]^2 = [a(a - 1)(a + 1)(a + 2)]^2/144$$

Soit finalement :

$$\sum_{i=1,a} [E(t_i - E(t_i))^2]^2 = a(a^2 - 1)(a + 2)(2a + 1)(5a^2 + 5a - 9)/630$$

Le moment centré d'ordre 4 s'obtient en utilisant la relation (3) dont tous les éléments sont connus. ♦

Nous avons jusqu'à présent utilisé la proposition 1 pour obtenir explicitement les 4 premiers moments de θ_a . La proposition 1 peut également être exploitée pour obtenir la fonction génératrice. Plus précisément, on obtient une relation de récurrence pour la fonction génératrice de θ_a :

PROPOSITION 2 : La fonction génératrice $h_i(s)$ de t_i , vérifie la relation de récurrence suivante :

$$h_i(s) = s/[2 - sh_{i-2}(s)]$$

De plus on a $h_0(s) = 1$ et $h_1(s) = s$.

Démonstration : Il suffit d'exploiter la relation de récurrence de la proposition 1 pour obtenir

$$h_i(s) = s(1 + h_i(s)h_{i-2}(s))/2 \quad \blacklozenge$$

Le corollaire 3 permet d'obtenir la fonction génératrice $g_a(s)$ de θ_a .

On obtient avec un logiciel de calcul formel $g_a(s)$ pour n'importe quelle valeur de a ; par exemple :

$$g_1(s) = s$$

$$g_2(s) = s^2/(2 - s)$$

$$g_3(s) = s^3/[(s - 2)(s^2 - 2)] \dots$$

$$g_8(s) = s^8/[(8 - 8s + s^4)(16 - 8s - 12s^2 + 4s^3 + s^4)]$$

$$g_9(s) = s^9/[(16 - 8s - 12s^2 + 4s^3 + s^4)(16 - 20s^2 + 5s^4)]$$

On peut ainsi obtenir n'importe quel moment en dérivant la fonction génératrice ⁽¹⁾.

REMERCIEMENTS

Je remercie un lecteur anonyme dont les remarques ont permis de rendre plus claire la démonstration de la proposition 1.

RÉFÉRENCES

- P. VALLOIS et C. S. TAPIERO, Moments of an Amplitude Process in a Random Walk and Approximations: Computations and Applications, à paraître dans *Recherche opérationnelle*.
- P. VALLOIS, On the range process of a bernoulli random walk, *Proceedings of the sixth International Symposium on Applied Stochastic Models and Data Analysis*, Vol. II. J. Janssen and C. H. Skiadas ed., World Scientific, p. 1020-1031.

⁽¹⁾ Les programmes écrits en Maple sont disponibles auprès de l'auteur.